Engineering Letters, 13:3, EL_13_3_5 (Advance online publication: 4 November 2006)

Automatic Building Extraction from Satellite Imagery

Lau Bee Theng, *Member, IAENG*

Abstract – **Automatic building extraction is an active research in remote sensing recently. It has been going on for more than 20 years but the automated extractions still encounter problems due to image resolution, variation and level of details. Mayunga et al. (2005) developed an improved snake model. However their radial casting encounters difficulties in initializing the snake model. This research paper discusses the development of an active contour model initialization algorithm. The prototype uses the existing improved snake energy function to compute the snake contours, but initialize the model with circular casting algorithm instead of radial casting algorithm.**

I. Introduction

Most of the building extraction methods use generic models by assuming all buildings follow a certain pattern. Hence the generic models do not provide practical results when buildings in unstructured buildings like informal settlements precede the particular pattern (Ruther et al. 2002). Buildings in informal settlements areas differ in terms of building materials, neighborhood distance and orientations. There are limited tools and methods to extract unstructured buildings as compared to researches on structured building extraction. Recently some researches are published such as use of fused shadow data with 2D building blobs derived from normalized Digital Surface Model (DSM) (Li and Ruther 1999) and still video Kodak camera to extract shacks in South Africa (Baltsavias and Mason 1995). However DSM suffers from insufficient ground sampling data and matching errors due to poor image quality, and also occlusion and shadows that lead to poor definition of buildings outlines (Baltsavias and Mason 1995 and Ruther et al. 2002). An effective informal settlements extractor should accommodate both structured and unstructured buildings (Ruther et al. 2002).

Generally, the main tasks in building extraction from digital images are building detection and building reconstruction. However building extraction tasks may differ depending on the use of geometrical representation with rectangular models (Weidner and Frostner 1995), use of multiple images (Ballard and Zisserman 2000), and polyhedral shapes (Scholtze et al. 2002), the use of lines, points and regions to describe building outlines (Fisher et al. 1997). The existing automated building extraction techniques are still performing at elementary level caused by image variation in terms of type, scale, and required level of detail (Wang and Tseng 2003). Furthermore automatic recognition of semantic information of an object using computers is complicated; most algorithms fail whenever a new situation in image space is encountered or objects are close to each other (Gruen 2000, Ruther et al. 2002). Sohn and Dowman (2001) proposed an automatic method of extracting buildings in densely urban areas from IKONOS images. They used large detached buildings without analysis of accuracy and structure details. Ortner (2002) implemented optimization and destruction approaches concurrently for building extraction. A point process technique is developed to extract well-structured buildings. Fraser et al. (2002) compared buildings extracted from IKONOS imagery with those obtained using black and white aerial photographs to evaluate the potential of high-resolution images. Toutin and Cheng (2002) investigated the potential of Quickbird imagery for spatial data acquisition, they found out that sensors of 0.6m have lessened the gap between satellite images and aerial photographs that have resolution from 0.2 to 0.3m. Haverkamp (2003) implemented a linking edge chain to extract buildings from IKONOS images. Thomas et al (2003) concluded that high-resolution imagery is a valuable tool for mapping urban areas in extracting land cover information from high-resolution images.

II. Active Contour Model

Active contour model is a useful model to extract structured and unstructured objects from digital images for informal settlements (Mayunga et al. 2005). It is defined by an energy function. The energy function (E_{Snakes}) is a weighted combination of internal (E_{Inter}), image (E_{Img}) and external constraint (E_{Cont}) forces as in Equation 1. The internal energy force describes the shape of the active contour; the image force attracts the active contours to the boundaries of the object, the external energy force comes from the image itself or higher level image processing. The solution of the active contour model is activated by its intrinsic trend of minimizing its energies. The energy function reaches minimum when the active contours control points locks the object boundaries in the image space. Kaas et al (1988) represented a contour by a vector, V(s) = [x(s), y(s)], having the arc length *s* as a parameter, where *x* and *y* are the coordinates of a active contour point. The total energy of an active contour is:

$$E_{Snakes} = \int_0^1 E_{Inter}V(s) + E_{Img}V(s) + E_{Cont}V(s)ds \qquad (1)$$

Several approaches have been proposed to remedy the above-mentioned problems. For example; the use of active contours and least squares method to extract buildings in 2D and 3D using aerial photography and satellite images (Mayunga et al. 2005). Cohen and Cohen (1990) used pressure force to control the movement of active contour. Although the method worked well, the parameter that controls the inflating force is difficult to estimate for high level noise in imagery. Tabb et al. (2000) combined active contours and neural networks to detect and categorize objects in images. The active contour is stored as a vector of (x, y) coordinates reflecting the position of different control point on the contour's spline. The coordinates are the input for neural network. Kreschner (2001) used homologous twin active contours and integrated in a bundle adjustment. The method fails when the system chooses wrong active contour. Ruther et al. (2002) use active contour and dynamic programming optimization technique to model buildings in informal settlement areas. However, dynamic programming is computational expensive and fails in more complex topologies. Guo and Yasuoka (2003) adopted a "balloon snake model". Multiple Height Bin (MHB) technique was employed to obtain the approximate active contours. MHB technique could not provide correct representation of the extracted objects.

In Mayunga et al. (2005) proposed an improved active contour energy function with radial casting initialization, the external energy that creates boundary effects for unstructured buildings and weighted coefficients are discarded. The algorithm being minimized is expressed as:

$$\begin{aligned} E_{Snakes} &= E_{\mathrm{Im}g} + E_{Cont} + E_{Curv} \\ E_{\mathrm{Im}g} &= \frac{(Min - Mag)}{(Max - Min)} \\ E_{Cont} &= \sum_{i=1}^{n} \frac{|v_{i+1} - v_i|}{n} - |v_i - v_{i-1}| \\ E_{Curv} &= |v_{i-1} - 2v_i + v_{i+1}|^2 \\ v_i &= (x_i, y_i) \end{aligned} \qquad (2)$$

It consists of three energy terms that are Continuity, Curvature term and Image. Image term describes the radiometric content of the image and it restricts the active contour points to move towards the points of highest gradient. The gradient of image at each control point is normalized to show small differences in values at the neighborhood of that control point. In this case, the gradient magnitude is negative to enable control points with large gradient to have small values. The image term attract the contours to the image points with minimum gradient magnitude. Continuity term creates equal space contours control points to avoid grouping and minimize distances. Curvature term expresses the curvature of the active contours.

III. Active contour Initialization

Mayunga et al (2005) proposed a radial casting algorithm for initiating the active contour model as shown in Figure 1. The contour's centre point C is measured, and from this point radial lines are projected outwards at definable angular intervals, β consists of four, eight, or sixteen radial lines ranging from 0 to 360. Each active contour control point in image space, advance to a new position where the gradient energy in a search window is maximum. The advancement has to proceed using angular intervals as well. The centre point C of the building polygon is always fixed and the radial distances, *l* to the active contour control points is variable depending on the size of the building object.

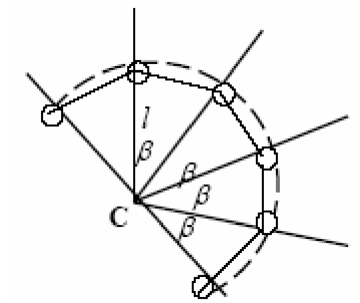

Figure 1 Radial Cast Model

The problem with this radial case is the number of radial lines depends on the complexity of the building. Hence more complex structure would require a high volume of lines to make the active contours. Furthermore the radial cast model claimed to be effective where a maximum of 16 radical lines should be sufficient to extract a building. However this is not feasible where unstructured building has irregular shape as shown in Figure 2. The extracted building C varied from the original building A, object B shows the 16 radial lines derived for the building from its centre.

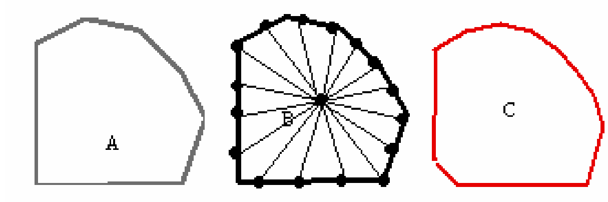

Figure 2 Irregular Shape and Radial Casting

On the other hand, the active contour control points from point C possibly cause the curve to be smaller than desired during radiation. If it happened, the generated active contour control points are deleted and a new centre point C is established. This causes inconsistency for all other control points built unless they are all deleted and rebuilt.

IV. Proposed Active contour Initialization

As response to the problems in existing radial casting, the circular cast initialization for active contour is developed. A control point, *C* can be measured at any point of the building object not necessarily the center as picking up center requires manual operator. Hence this allows automation where the point can be picked by comparing pixel using corner detector (Haris and Stephens 1988). Corner point detector works by searching for pixel in an object based on the assumption that corners are associated with maxima of the local autocorrelation function. It calculates corner value for each pixel of the image, and if *C* is a local maximum above certain threshold, the pixel is declared a corner. Consequently a circular cast is initiated from any of this first found corner pixel as control point. The circular cast has a diameter of 37 pixels as it is the standard use in image edge and feature detection (Perez and Dennis 1997).

```
--------------------------------------------------------------
Set local maxima = control point
while extraction is not closed
   create circular cast
   while there is more pixel  within circular cast
            get pixel
            if pixel attribute = local maxima
                append pixel as a node to active contour
            end if
 end while
end while
--------------------------------------------------------------
```

Figure 3 Circular Cast Algorithm

The circular cast built for the irregular shape is similar to building A in Figure 2. The circular cast resolves the inability of radial cast to cope for irregular shape as proposed in Mayunga et al. (2005). The algorithm works by checking all the pixels with the cast's boundary, the extracted pixel is compared to the local maxima obtained from the control point, the cast stops expanding when the extraction is closed. In Figure 4 illustration A to D shows the inflation of the cast as the contour is incomplete, illustration E shows the completed extraction where the contour is fully closed.

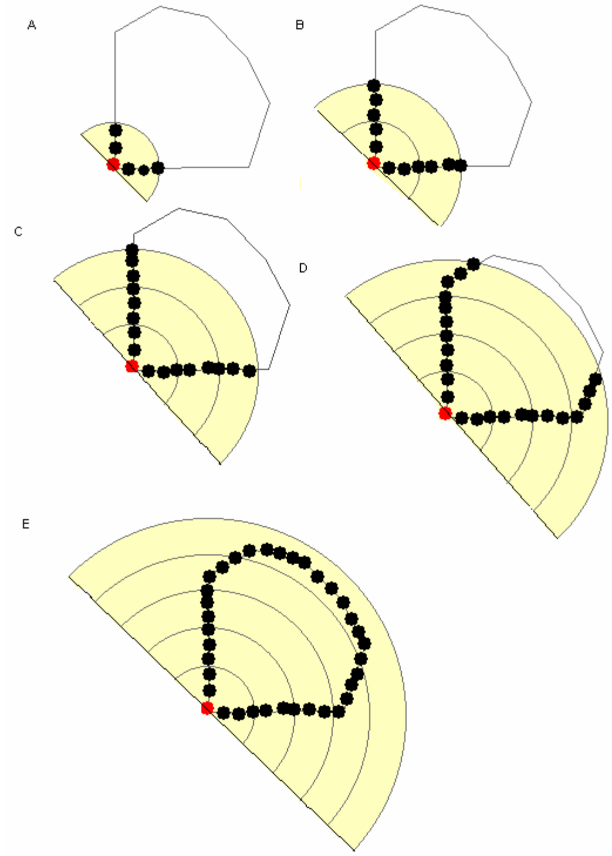

Figure 4 Sample of Circular Cast built for Irregular Building Structure from Step A to E

V. PROTOTYPING

The proposed active contour initialization methodology for buildings extraction from satellite imagery is prototyped. The prototype consists of image pre-processing and building extraction. The variation in illumination conditions, shadowing and building density in urban areas makes it very difficult to distinguish individual buildings from its surrounding. In order to solve this problem, a non-linear anisotropic diffusion model (Weickert 1999) was used to normalize the noise effects around the buildings. The image normalization process brings the variation of pixels around the buildings at the same level. The diffused image is then used as an input for building extraction.

Once a corner is obtained for a building in the image space and then the active contour are automatically generated. User is given an option to accept or reject a single snake contour or all generated snake contours. If a snake contour is accepted, an iterative minimization function is invoked, minimum and maximum energy values in the neighborhood are computed. Neighborhood point with the lowest energy value is the new position in the image space. The iteration stops when active contour locks a building outline.

The proposed method was applied to extract buildings from the buildings at Burlington city. The images used were obtained from the website and geo-rectified by the image vendor DigitalGlobe, it has 4-band spectral resolution at 16 bits/pixel, pre-processed by vendor and geometrically corrected. The prototype is implemented using a desktop (Pentium (R) 4 CPU 2.99 GHZ processor), VBA and ArcScene. Figure 5 show a portion of Quickbird imagery for Burlington, Figure 6 shows the test area, Figure 7 shows extracted buildings. Figure 8 shows 2D buildings in vectorized layer.

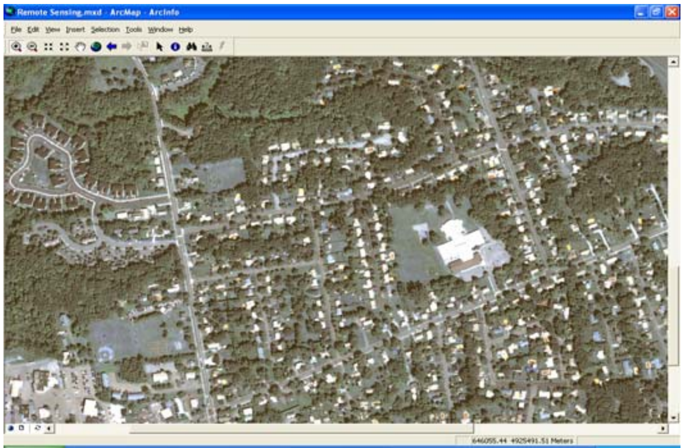

Figure 5 A Quickbird Imagery of Burlington

Figure 6 Imagery for the Test Area

Figure 7 Extracted Buildings from the Test Area

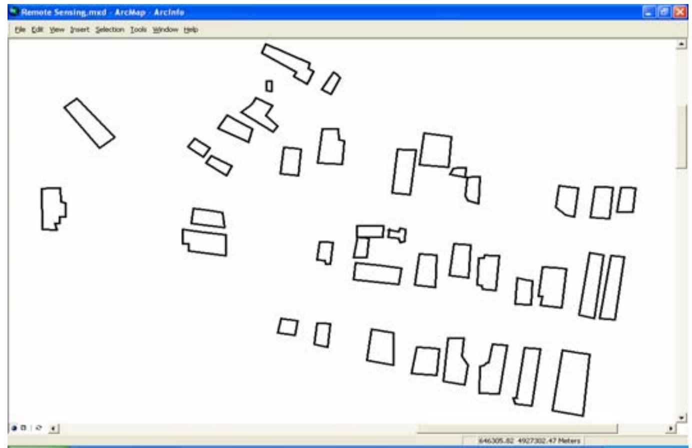

Figure 8 Extracted Area in Vectorized View

VI. QUANTITATIVE ANALYSIS

A total of 100 building corner points from 2D vectorized layer were randomly compared with their corresponding points from the ground truth data.

No. of points	RMSE (m)	Std dev. in x	Std dev. in y
100	**1.0**	**0.75**	**0.95**

Figure 9 The RMSE and deviations of randomly measured building corner points

Root Mean Square Error (RMSE) was computed to determine the internal accuracy of the measurement. Standard deviations in (x, y) were also computed in Table 2. The possible reason contributing to this figure is closeness of buildings and resolution of the image. Informal settlements consists of random noise causes edges along the corners to divert from their correct positions. There is a need for post-processing stage to refine the edges if higher accuracy application is required.

CONCLUSION

The improved snake energy function (Mayunga 2005) works well with proposed circular casting algorithm in this study. It improved the RMS and standard deviations from the ground truth data. It is a significant contribution to building extraction from high-resolution satellite imagery. The approach has been tested on structured and unstructured buildings in a Burlington urban settlement. Buildings with irregular shapes are extracted with reliable accuracy.

REFERENCES

[1] Ballard, C. and Zisserman, A. (2000). A Plane-sweep Strategy for the 3D Reconstruction of Buildings from Multiple Images. *IASPRS*, Vol. XXXIII, Part B2, pp 56-62

[2] Baltsavias, E. and Mason, S. (1995). *Image–Based Reconstruction of Informal Settlements. Automatic Extraction of Man Made Objects from Aerial and Space Images II*, Birkhauser Verlag, Berlin, Germany, pp. 97-108.

[3] Cohen, L. D. and Cohen, I. (1990). A Finite Element Method Applied to New Active Contour Models and 3-D Reconstruction from Crossing Sections. In *Proceedings of the Third International Conference on Computer Vision*, Osaka, pp. 587-591.

[4] Digital Globe Inc. QuickBird Sample Imagery. Accessed 30/03/06. http://www.digitalglobe.com/sample_imagery.shtml.

[5] Fraser C. S., Baltsavias, E. and Gruen A. (2002). Processing of IKONOS Imagery for Submeter 3D Positioning and Building Extraction. *International Journal of Photogrammetry and Remote Sensing* Vol. 56 pp. 177-194.

[6] Fisher, A., Kolbe, T. H. and Lang, F. (1997). Integration of 2D and 3D Reasoning for Building Reconstruction Using a Generic Hierarchical Model, In *Proceedings of the Workshop on Semantic Modeling for the Acquisition of Topographic Information from Images and Maps SMATI'97*, pp. 159-180.

[7] Guo, T. and Yasuoka, Y. (2003). Snake-based Approach for Building Extraction from High-Resolution Satellite Images and Height Data in Urban Areas. Accessed 30/03/06. http://www.gisdevelopment.net/aars/acrs/2002/vhr/018.pdf

[8] Gruen, A. (2000). Potential and Limitation of High-Resolution Satellite Imagery. *Proceedings of the Asian Conference on Remote Sensing,* December 4-8, Taiwan.

[9] Haris, C. and Stephens, M. (1988). A Combined Corner and Edge Detector. In *Proceedings of 4th Alvey Vision Conference*, Manchester, pp. 141-151.

[10] Haverkamp, D. (2003). Automatic Building Extraction from IKONOS Imagery, Accessed 30/03/06. http://www.spaceimaging.com/whitepapers_pdfs /2004/

[11] Kaas, M., Witkin, A. and. Terzopoulos, D. (1988). Snakes: Active Contour Models. *International Journal of Computer Vision*, (1): p321-331.

[12] Kreschner, .M. (2001). Homologous Twin Snakes Integrated in a Bundle Block Adjustment. Accessed 30/03/06. http://www.ipf.tuwien.ac.at/publications/mk_p_isprs98/paper3.html

[13] Lau, B. T. (2006). Automatic Informal Settlements Building Extraction from Satellite Imagery. In the Proceedings of *3rd International Conferences on Artificial Intelligence in Engineering and Technology (ICAIET06).* Kota Kinabalu.

[14] Li, J. and Ruther, S. H. (1999). IS-Modeller: A low-Cost Image-based Tool for Informal Settlements Planning, *Geoinfomatics '99 Conference*, Ann Arbor, 19-21 June, 1999, pp. 1-9.

[15] Mayunga, S. D., Zhang, Y. and Coleman, D. J. (2005). Semi-Automatic Building Extraction Utilizing Quickbird Imagery. In *Stilla, U., Rottensteiner, F. and Hinz, S. (Eds.), IAPRS*, Vol. XXXVI, Part 3/W24, pp. 131-136.

[16] Otner, M., Descombes, X. and Zerubia J. (2002). Building Extraction from Digital Elevation Model, *Rapport INRIA No. 4517*, INRIA France.

[17] Perez, M. M. and Dennis, T. J. (1997). Adaptive Implementation of the SUSAN Method for Image Edge and Feature Detection. In *Proceedings of IEEE Conference on Image Processing*, pp. 394-397.

[18] Ruther, H., Hagai, M. M and Mtalo, E.G. (2002). Application of Snakes and Dynamic Programming Optimization in Modeling of Buildings in Informal Settlement Areas, *ISPRS Journal of Photogrammetry and Remote Sensing*, 56(2002): p 269-282.

[19] Scholze S., T., Moons, T. and Gool, V. (2002). A Probabilistic Approach to Roof Extraction and Reconstruction. *IASPRS*, Commission III, pp. 231-237.

[20] Sohn, G. and Dowman, I. (2001). Extraction of Buildings from High Resolution Satellite Data. In *Baltsavias, E., Gruen, A., Van Gool, L. (Eds.), Automated Extraction of Man-Made Objects from Aerial and Space Images (III).* Balkema Publishers, Lisse, pp. 345-355.

[21] Tabb, K., Davey, N. and Stella, G. (2000). Analysis of Human Motion using Snakes and Neural Networks. Accessed 30/03/06. http://homepages.feis.herts.ac.uk/~nngroup/pubs/papers/tabb-amdo00.pdf

[22] Toutin, T. and Cheng, P. (2002). 3D Models for High Resolution Images: Examples with QuickBird, IKONOS and EROS. *Symposium on Geospatial Theory, Processing and Applications*, Ottawa.

[23] Thomas, N., Hendrix, C. and Cogalton, R.G. (2003). A Comparison of Urban Mapping Methods Using High-Resolution Digital Imagery. *Photogrammetric Engineering and Remote Sensing,* Vol. 69, No. 9, September 2003, pp. 963-972.

[24] Weidner, U. and Forstner, W. (1995). Towards Automatic Building Extraction from High Resolution Digital Elevation Models. *International Society of Photogrammetry and Remote Sensing,* Vol. 50 No. 4, pp 38-49.

[25] Wang, S. and Tseng, Y.H. (2003). Semi-automated Building Extraction Based on CSG Model-Image Fitting. *Photogrammetry Engineering & Remote Sensing* Vol. 69, No. 2: pp. 171-180

[26] Weickert, J. (1999). Coherence-enhancing Diffusion Filtering. *International Journal of Computer Vision*, Vol. 31 pp. 111-127.